

FIGURE 1

CGGACGCGTGGGACCCATACTTGCTGGTCTGATCCATGCACAAGGCGGGGCTGCTAGGCCTC
TGTGCCCCGGGCTTGGAATTCGGTGCGGATGGCCAGCTCCGGGATGACCCGCCGGGACCCGCT
CGCAAATAAGGTGGCCCTGGTAACGGCCTCCACCGACGGGATCGGCTTCGCCATCGCCCGGC
GTTTGGCCCAGGACGGGGCCCATGTGGTCGTCAGCAGCCGGAAGCAGCAGAATGTGGACCAG
GCGGTGGCCACGCTGCAGGGGGAGGGGCTGAGCGTGACGGGCACCGTGTGCCATGTGGGGAA
GGCGGAGGACCGGGAGCGGCTGGTGGCCACGGCTGTGAAGCTTCATGGAGGTATCGATATCC
TAGTCTCCAATGCTGCTGTCAACCCTTTCTTTGGAAGCATAATGGATGTCACTGAGGAGGTG
TGGGACAAGACTCTGGACATTAATGTGAAGGCCCCAGCCCTGATGACAAAGGCAGTGGTGCC
AGAAATGGAGAAACGAGGAGGCGGCTCAGTGGTGATCGTGTCTTCCATAGCAGCCTTCAGTC
CATCTCCTGGCTTCAGTCCTTACAATGTCAGTAAAACAGCCTTGCTGGGCCTGACCAAGACC
CTGGCCATAGAGCTGGCCCCAAGGAACATTAGGGTGAAGTGCCTAGCACCTGGACTTATCAA
GACTAGCTTCAGCAGGATGCTCTGGATGGACAAGGAAAAAGAGGAAAGCATGAAAGAAACCC
TGCGGATAAGAAGGTTAGGCGAGCCAGAGGATTGTGCTGGCATCGTGTCTTTCCTGTGCTCT
GAAGATGCCAGCTACATCACTGGGGAAACAGTGGTGGTGGGTGGAGGAACCCCGTCCCGCCT
CTGAGGACCGGGAGACAGCCCACAGGCCAGAGTTGGGCTCTAGCTCCTGGTGCTGTTTCCTGC
ATTCACCCACTGGCCTTTCCACCTCTGCTCACCTTACTGTTACCTCATCAAATCAGTTCT
GCCCTGTGAAAAGATCCAGCCTTCCTGCCGTCAAGGTGGCGTCTTACTCGGGATTCTTGCT
GTTGTTGTGGCCTTGGGTAAAGGCCTCCCTGAGAACACAGGACAGGCCTGCTGACAAGGCT
GAGTCTACCTTGGCAAAGACCAAGATATTTTTTCTGGGCCACTGGTGAATCTGAGGGGTGA
TGGGAGAGAAGGAACCTGGAGTGGAAGGAGCAGAGTTGCAAATTAACAGCTTGCAAATGAGG
TGCAAATAAAATGCAGATGATTGCGCGGCTTTGAAAAAAAAA

FIGURE 2

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA35672

><subunit 1 of 1, 278 aa, 1 stop

><MW: 29537, pI: 8.97, NX(S/T): 1

MHKAGLLGLCARAWNSVRMASSGMTRRDPLANKVALVTASTDGIGFAIARRLAQDGAHVVS
SRKQQNVNQAVATLQGEGLSVTGTVCHVGKAEDRERLVATAVKLHGGIDILVSNAAVNPFFG
SIMDVTEEVWDKTLTDINVKAPALMTKAVVPEMEKRGGGSVVIVSSIAAFSPSPGFSPYNVSK
TALLGLTKTLAIELAPRNIRVNCLAPGLIKTSFSRMLWMDKEKEESMKETLRIRRLGEPEDC
AGIVSFLCSEDASYITGETVVVGGGTPSRL

Important features of the protein:

Signal peptide:

amino acids 1-15

N-glycosylation site.

amino acids 183-186

N-myristoylation sites.

amino acids 43-48, 80-85, 191-196, 213-218, 272-277

Microbodies C-terminal targeting signal.

amino acids 276-278

bioRxiv preprint doi: <https://doi.org/10.1101/066201>; this version posted November 1, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

FIGURE 3

GCGCCCTGAGCTCCGCCTCCGGGCCCGATAGCGGCATCGAGAGCGCCTCCGTCGAGGACCAGGCGGCG
CAGGGGGCCGGCGGGCGAAAGGAGGATGAGGGGGCGCAGCAGCTGCTGACCCTGCAGAACCAGGTGGC
GCGGCTGGAGGAGGAGAACCAGACTTTCTGGCTGCGCTGGAGGACGCCATGGAGCAGTACAACTGC
AGAGCGACCGGCTGCGTGAGCAGCAGGAGGAGATGGTGGAACTGCGGCTGCGGTTAGAGCTGGTGC
CCAGGCTGGGGGGCCTGCGGCTCCTGAATGGCCTGCCTCCCGGGTCCTTTGTGCCTCGACCTCATAC
AGCCCCCTGGGGGGTGCCACGCCCATGTGCTGGGCATGGTGCCGCTGCCTGCCTCCCTGGAGATG
AAGTTGGCTCTGAGCAGAGGGGAGAGCAGGTGACAAATGGCAGGGAGGCTGGAGCTGAGTTGCTGACT
GAGGTGAACAGGCTGGGAAGTGGCTCTTCAGCTGCTTCAGAGGAGGAAGAGGAGGAGGAGGAGCCGCC
CAGGCGGACCTTACACCTGCGCAGAAATAGGATCAGCAACTGCAGTCAGAGGGCGGGGGCACGCCAG
GGAGTCTGCCAGAGAGGAAGGGCCAGAGCTTTGCCTTGAGGAGTTGGATGCAGCCATTCCAGGGTCC
AGAGCAGTTGGTGGGAGCAAGGCCCGAGTTCAGGCCCGCCAGGTCCCCCTGCCACAGCCTCAGAGTG
GCGGCTGGCCCAGGCCAGCAGAAGATCCGGGAGCTGGCTATCAACATCCGCATGAAGGAGGAGCTTA
TTGGCGAGCTGGTCCGCACAGGAAGGCAGCTCAGGCCCTGAACCGCCAGCACAGCCAGCGTATCCGG
GAGCTGGAGCAGGAGGCAGAGCAGGTGCGGGCCGAGCTGAGTGAAGGCCAGAGGCAGCTGCGGGAGCT
CGAGGGCAAGGAGCTCCAGGATGCTGGCGAGCGGTCTCGGCTCCAGGAGTTCGCGAGGAGGGTCGCTG
CGGCCAGAGCCAGGTGCAGGTGCTGAAGGAGAAGAAGCAGGCTACGGAGCGGCTGGTGTCACTGTGCG
GCCAGAGTGAGAAGCGACTGCAGGAGCTCGAGCGGAACGTGCAGCTCATGCGGCAGCAGCAGGGACA
GCTGCAGAGGCGGCTTCGCGAGGAGACGGAGCAGAAGCGGCGCCTGGAGGCAGAAATGAGCAAGCGGC
AGCACCGCGTCAAGGAGCTGGAGCTGAAGCATGAGCAACAGCAGAAGATCCTGAAGATTAAGACGGAA
GAGATCGCGGCTTCCAGAGGAAGAGGCGCAGTGCCAGCAACGGCTCTGTGGTCAGCCTGGAACAGCA
GCAGAAGATTGAGGAGCAGAAGAAGTGGCTGGACCAGGAGATGGAGAAGGTGCTACAGCAGCGGCGGG
CGCTGGAGGAGCTGGGGGAGGAGCTCCACAAGCGGGAGGCCATCCTGGCCAAGAAGGAGGCCCTGATG
CAGGAGAAGACGGGGCTGGAGAGCAAGCGCTGAGATCCAGCCAGGCCCTCAACGAGGACATCGTGCG
AGTGTCCAGCCGGCTGGAGCACCTGGAGAAGGAGCTGTCCGAGAAGAGCGGGCAGCTGCGGCAGGGCA
GCGCCAGAGCCAGCAGCAGATCCGCGGGGAGATCGACAGCCTGCGCCAGGAGAAGGACTCGTGCTC
AAGCAGCGCCTGGAGATCGACGGCAAGCTGAGGCAGGGGAGTCTGCTGTCCCCGAGGAGGAGCGGAC
GCTGTTCCAGTTGGATGAGGCCATCGAGGCCCTGGATGCTGCCATTGAGTATAAGAATGAGGCCATCA
CATGCCGCCAGCGGGTGCTTCGGGGCTCAGCCTCGTTGCTGTCCAGTGCGAGATGAACCTCATGGCC
AAGCTCAGCTACCTCTCATCCTCAGAGACCAGAGCCCTCCTCTGCAAGTATTTTGACAAGGTGGTGAC
GCTCCGAGAGGAGCAGCACCCAGCAGCAGATTGCCTTCTCGGAACTGGAGATGCAGCTGGAGGAGCAGC
AGAGGCTGGTGTACTGGCTGGAGGTGGCCCTGGAGCGGCAGCGCCTGGAGATGGACCGCCAGCTGACC
CTGCAGCAGAAGGAGCACGAGCAGAACATGCAGCTGCTCCTGCAGCAGAGTCGAGACCACCTCGGTGA
AGGGTTAGCAGACAGCAGGAGGCAGTATGAGGCCCGGATTCAAGCTCTGGAGAAGGAACTGGGCGGTT
ACATGTGGATAAACCAGGAACCTGAAACAGAAGCTCGGCGGTGTGAACGCTGTAGGCCACAGCAGGGGT
GGGAGAAGAGGAGCCTGTGCTCGGAGGGCAGACAGGCTCCTGGAAATGAAGATGAGCTCCACCTGGC
ACCCGAGCTTCTCTGGCTGTCCCCCTCACTGAGGGGGCCCCCGCACCCGGGAGGAGACGCGGGACT
TGGTCCACGCTCCGTTACCCTTGACCTGGAACGCTCGAGCCTGTGTGGTGAGGAGCAGGGGTCCCC
GAGGAACCTGAGGCAGCGGGAGGCGGCTGAGCCCCCTGGTGGGGCGGGTGCTTCTGTGGGTGAGGCAGG
CCTGCCCTGGAACCTTGGGCCTTTGTCCAAGCCCCGGCGGGAACCTGCGACGAGCCAGCCCGGGGATGA
TTGATGTCCGGAAAAACCCCTGTAAAGCCCTCGGGGCAGACCCTGCCTTGGAGGGAGACTCCGAGCCT
GCTGAAAGGGGCAGCTGCCTGTTTGTCTTCTGTGAAGGGCAGTCCTTACCGCACACCCTAAATCCAGG
CCCTCATCTGTACCTCACTGGGATCAACAAATTTGGGCCATGGCCAAAAGAACTGGACCCTCATTT
AACAAAATAATATGCAATTCCCACTTACTTCCATGAAGCTGTGGTACCCAATTGCCGCCTTGTG
TCTTGCTCGAATCTCAGGACAAATCTGGTTTTAGGCGTAAATGGATGTGCTTGTAGTTTCAAGGGTTT
GCCAAGAATCATCACGAAAGGGTCGGTGGCAACCAGGTTGTGGTTTAAATGGTCTTATGTATATAGGG
GAAACTGGGAGACTTTAGGATCTTAAAAAACCATTTAATAAAAAAAATCTTTGAAGGGAC

FIGURE 4

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA47465

<subunit 1 of 1, 830 aa, 1 stop

<MW: 95029, pI: 8.26, NX(S/T): 2

MEQYKLQSDRLREQQEEMVELRLRLELVRPGWGGLRLLNGLPPGSFVPRPHTAPLGGAHAHV
LGMVPPACLPGDEVGSEQRGEQVTNGREAGAELLTEVNRLGSGSSAASEEEEEEEPPRRTL
HLRRNRISNCSQRAGARPGSLPERKGPELCEELDAAIPGSRAVGGSKARVQARQVPPATAS
EWRLAQAQQKIRELAINIRMKEELIGELVRTGKAAQALNRQHSQRIRELEQEAEQVRAELSE
GQRQLRELEGKELQDAGERSRLQEFRRRVAAAQSQVQVLKEKKQATERLVSLSAQSEKRLQE
LERNVQLMRQQQGQLQRRRLREETEQRRLLEAEMSKRQHRVKELELKHEQQQKILKIKTEEIA
AFQRKRRSGSNGSVVSLEQQQKIEEQKKWLDQEMEKVLQQRRALEELGEELHKREAILAKKE
ALMQEKTGLESKRLRSSQALNEDIVRVSSRLEHLEKELSEKSGQLRQGSQAQSQQQIRGEIDS
LRQEKDSLKQRLQLEIDGKLQGSLLSPEEERTLFQLDEAIEALDAAIEYKNEAITCRQRVLR
ASASLLSQCEMNLMAKLSYLSSSETRALLCKYFDKVVTLREEQHQQQIAFSELEMQLEEQQR
LVYWLEVALERQRLEMDRQLTLQQKEHEQNMQLLLQQSRDHLGEGGLADSRQYEARIQALEK
ELGRYMWINQELKQKLGGVNAVGHSRGGEKRSKSLCSEGRQAPGNEDELHLAPELLWLSPLTEG
APRTREETRDLVHAPLPLTWKRSSLCGEEQGSPEELRQREAAEPLVGRVLPVGEAGLPWNFG
PLSKPRRELRRASPGMIDVRKNPL

Important features:

Leucine zipper pattern.

amino acids 557-579, 794-815

N-glycosylation sites.

amino acids 133-136, 383-386

Kinesin related protein Kif-4 Coiled-coil domain:

amino acids 231-672

2.4	
3.1	1.00
3.2	1.00
3.3	1.00
3.4	1.00
3.5	1.00
3.6	1.00
3.7	1.00
3.8	1.00
3.9	1.00
4.0	1.00
4.1	1.00
4.2	1.00
4.3	1.00
4.4	1.00
4.5	1.00
4.6	1.00
4.7	1.00
4.8	1.00
4.9	1.00
5.0	1.00
5.1	1.00
5.2	1.00
5.3	1.00
5.4	1.00
5.5	1.00
5.6	1.00
5.7	1.00
5.8	1.00
5.9	1.00
6.0	1.00
6.1	1.00
6.2	1.00
6.3	1.00
6.4	1.00
6.5	1.00
6.6	1.00
6.7	1.00
6.8	1.00
6.9	1.00
7.0	1.00
7.1	1.00
7.2	1.00
7.3	1.00
7.4	1.00
7.5	1.00
7.6	1.00
7.7	1.00
7.8	1.00
7.9	1.00
8.0	1.00
8.1	1.00
8.2	1.00
8.3	1.00
8.4	1.00
8.5	1.00
8.6	1.00
8.7	1.00
8.8	1.00
8.9	1.00
9.0	1.00
9.1	1.00
9.2	1.00
9.3	1.00
9.4	1.00
9.5	1.00
9.6	1.00
9.7	1.00
9.8	1.00
9.9	1.00
10.0	1.00

Variable	Mean	SD	Min	Max
Age	35.2	10.5	21	55
Gender	1.2	0.4	1	2
Marital Status	1.5	0.5	1	3
Education	12.5	1.5	9	16
Income	3500	1500	1000	8000
Health Status	1.8	0.6	1	3
Stress Level	2.5	0.8	1	4
Life Satisfaction	3.2	0.9	1	5
Work-Life Balance	2.8	0.7	1	4
Family Support	3.5	0.8	1	5
Community Involvement	2.2	0.6	1	4
Personal Growth	3.0	0.7	1	5
Relationship Quality	3.8	0.9	1	5
Financial Stability	2.9	0.8	1	4
Healthcare Access	3.1	0.7	1	5
Work Environment	2.7	0.6	1	4
Education Quality	3.3	0.8	1	5
Income Stability	3.0	0.7	1	5
Health Insurance	3.4	0.8	1	5
Work-Life Balance	2.8	0.7	1	4
Family Support	3.5	0.8	1	5
Community Involvement	2.2	0.6	1	4
Personal Growth	3.0	0.7	1	5
Relationship Quality	3.8	0.9	1	5
Financial Stability	2.9	0.8	1	4
Healthcare Access	3.1	0.7	1	5
Work Environment	2.7	0.6	1	4
Education Quality	3.3	0.8	1	5
Income Stability	3.0	0.7	1	5
Health Insurance	3.4	0.8	1	5
Work-Life Balance	2.8	0.7	1	4
Family Support	3.5	0.8	1	5
Community Involvement	2.2	0.6	1	4
Personal Growth	3.0	0.7	1	5
Relationship Quality	3.8	0.9	1	5
Financial Stability	2.9	0.8	1	4
Healthcare Access	3.1	0.7	1	5
Work Environment	2.7	0.6	1	4
Education Quality	3.3	0.8	1	5
Income Stability	3.0	0.7	1	5
Health Insurance	3.4	0.8	1	5
Work-Life Balance	2.8	0.7	1	4
Family Support	3.5	0.8	1	5
Community Involvement	2.2	0.6	1	4
Personal Growth	3.0	0.7	1	5
Relationship Quality	3.8	0.9	1	5
Financial Stability	2.9	0.8	1	4
Healthcare Access	3.1	0.7	1	5
Work Environment	2.7	0.6	1	4
Education Quality	3.3	0.8	1	5
Income Stability	3.0	0.7	1	5
Health Insurance	3.4	0.8	1	5
Work-Life Balance	2.8	0.7	1	4
Family Support	3.5	0.8	1	5
Community Involvement	2.2	0.6	1	4
Personal Growth	3.0	0.7	1	5
Relationship Quality	3.8	0.9	1	5
Financial Stability	2.9	0.8	1	4
Healthcare Access	3.1	0.7	1	5
Work Environment	2.7	0.6	1	4
Education Quality	3.3	0.8	1	5
Income Stability	3.0	0.7	1	5
Health Insurance	3.4	0.8	1	5
Work-Life Balance	2.8	0.7	1	4
Family Support	3.5	0.8	1	5
Community Involvement	2.2	0.6	1	4
Personal Growth	3.0	0.7	1	5
Relationship Quality	3.8	0.9	1	5
Financial Stability	2.9	0.8	1	4
Healthcare Access	3.1	0.7	1	5
Work Environment	2.7	0.6	1	4
Education Quality	3.3	0.8	1	5
Income Stability	3.0	0.7	1	5
Health Insurance	3.4	0.8	1	5
Work-Life Balance	2.8	0.7	1	4
Family Support	3.5	0.8	1	5
Community Involvement	2.2	0.6	1	4
Personal Growth	3.0	0.7	1	5
Relationship Quality	3.8	0.9	1	5
Financial Stability	2.9	0.8	1	4
Healthcare Access	3.1	0.7	1	5
Work Environment	2.7	0.6	1	4
Education Quality	3.3	0.8	1	5
Income Stability	3.0	0.7	1	5
Health Insurance	3.4	0.8	1	5

FIGURE 6

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA57700

><subunit 1 of 1, 125 aa, 1 stop

><MW: 14198, pI: 9.01, NX(S/T): 1

MRPRCCILALVCWITVFLQCSKGTTDAPVGSGLWLCQPTPRCGNKIYNPSEQCCYDDAILS
LKETRRCGSTCTFWPCFELCCPESFGPQQKFLVKLRVLGMKSQCHLSPISRSCTRNRRLHVLYP

Important features:

Signal sequence

amino acids 1-21

N-myristoylation sites.

amino acids 33-39, 70-76

1003090123456789101112131415161718192021222324252627282930313233343536373839404142434445464748495051525354555657585960616263646566676869707172737475767778798081828384858687888990919293949596979899100

FIGURE 7

CCCACGCGTCCGCCCACGCGTCCGGGTGCCACTCGCGCGCCGGCCGCGCTCCGGGCTTCTCT
TTTCCCTCCGACGCGCCACGGCTGCCAGACATTCCGGCTGCCGGGTCTGGAGAGCTCCCCG
AACCCCTCCGCGGAGAGGAGCGAGGCGGCCAGGGTGGCCCCCGGGGCGCGCTTGGTCTCG
GAGAAGCGGGGACGAGGCCGGAGGATGAGCGACTGAGGGCGACGCGGGCACTGACGCGAGTT
GGGGCCGCGACTACCGGCAGCTGACAGCGCGATGAGCGACTCCCCAGAGACGCCCTAGCCCG
GTGTGCGCGCCAGGCGGAGCGCGCAGGTGGGGCTGGGCTGTTAGTGGTCCGCCCCACGCGGG
TCGCCGGCCGGCCCAGGATGGGCGCTGGCAACCCGGGCCCCGCGCCCGCTGCTACCCCTG
CGCCCGCTGCGAGCCCGGCGTCCGGCCCCGCGCCCTGCGCTCATGGACGGCGGGCTCCCGGCTG
GCGGCGGCGCGCCCCCGGGCTGTGAATGCGACTCGCCCCCTCGGCCGCGCTCCCCGCCCCGCC
GCCCCCGGGACGTGGTAGGGGATGCCCAGCTCCACTGCGATGGCAGTTGGCGCGCTCTCCA
GTTCCCTCCTGGTCACCTGCTGCCTGATGGTGGCTCTGTGCAGTCCGAGCATCCCGCTGGAG
AAGCTGGCCCAGGCACCAGAGCAGCCGGGGCCAGGAGAAGCGTGAGCACGCCACTCGGGACGG
CCCGGGGCGGGTGAACGAGCTCGGGCGCCCGGCGAGGGACGAGGGCGGCAGCGGCCGGGACT
GGAAGAGCAAGAGCGGCCGTGGGCTCGCCGGCCGTGAGCCGTGGAGCAAGCTGAAGCAGGCC
TGGGTCTCCCAGGGCGGGGGCGCCAAGGCCGGGGATCTGCAGGTCCGGCCCCGCGGGGACAC
CCCGCAGGCGGAAGCCCTGGCCGCAGCCGCCAGGACGCGATTGGCCCGGAACCTCGCGCCCA
CGCCCGAGCCACCCGAGGAGTACGTGTACCCGGAATAACCGTGGCAAGGGCTGCGTGGACGAG
AGCGGCTTCGTGTACGCGATCGGGGAGAAGTTTCGCGCCGGGCCCCCTCGGCCTGCCCGTGCCT
GTGCACCGAGGAGGGGCCGCTGTGCGCGCAGCCGAGTGCCCGAGGCTGCACCCGCGCTGCA
TCCACGTGACACGAGCCAGTGCTGCCCGCAGTGCAAGGAGAGGAAGAACTACTGCGAGTTC
CGGGGCAAGACCTATCAGACTTTTGAGGAGTTCGTGGTGTCTCCATGCGAGAGGTGTGCTG
TGAAGCCAACGGTGAGGTGCTATGCACAGTGTGAGCGTGTCCCCAGACGGAGTGTGTGGACC
CTGTGTACGAGCCTGATCAGTGCTGTCCCATCTGCAAAAATGGTCCAAACTGCTTTGCAGAA
ACCGCGGTGATCCCTGCTGGCAGAGAAGTGAAGACTGACGAGTGCACCATATGCCACTGTAC
TTATGAGGAAGGCACATGGAGAATCGAGCGGCAGGCCATGTGCACGAGACATGAATGCAGGC
AAATGTAGACGCTTCCCAGAACACAACTCTGACTTTTTCTAGAACATTTTACTGATGTGAA
CATTCTAGATGACTCTGGGAACATCAGTCAAAGAAGACTTTTGATGAGGAATAATGGAAAA
TTGTTGGTACTTTTTCTTTTCTTGATAACAGTTACTACAACAGAAGGAAATGGATATATTTT
AAAACATCAACAAGAACTTTGGGCATAAAATCCTTCTCTAAATAAATGTGCTATTTTACAG
TAAGTACACAAAAGTACACTATTATATATCAAATGTATTTCTATAATCCCTCCATTAGAGAG
CTTATATAAGTGTTTTCTATAGATGCAGATTAAAAATGCTGTGTTGTCAACCGTCAAAAAAA
AAAAAAAAAAAAAAAAAAAAA

FIGURE 8

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68818

><subunit 1 of 1, 325 aa, 1 stop

><MW: 35296, pI: 5.37, NX(S/T): 0

MPSSTAMAVGALSSSLLVTCCLMVALCSPSIPLEKLAQAPEQPGQEKREHATRDGPGRVNEL
GRPARDEGGSGRDWKSksGRGLAGREPWSKLKQAWVSQGGGAKAGDLQVRPRGDT PQAEALA
AAAQDAIGPELAPTPEPPPEEYVYPDYRGKGCVDES GFVYAIG EK FAPGPSACPCLCTEEGPL
CAQPEC PRLHPRCIHVDT SQCCPQCKERKNYCEFRGKTYQTLEEFVVSPCERCRC EANGEVL
CTVSACPQTECVDPVYEPDQCCPICKNGPNCFAETAVIPAGREVKTD ECTICHCTYE EGTWR
IERQAMCTRHECRQM

Important features of the protein:

Signal peptide:

amino acids 1-27

Transmembrane domain:

amino acids 11-30

Glycosaminoglycan attachment site.

amino acids 80-83

N-myristoylation sites.

amino acids 10-15, 102-107, 103-108

Cell attachment sequence.

amino acids 114-117

EGF-like domain cysteine pattern signature.

amino acids 176-187

1003290-12701
10/27/06 06:22:00

FIGURE 9

CAGCCACAGACGGGTCATGAGCGCGGTATTACTGCTGGCCCTCCTGGGGTTCATCCTCCAC
TGCCAGGAGTGCAGGCGCTGCTCTGCCAGTTTGGGACAGTTCAGCATGTGTGGAAGGTGTCC
GACCTACCCCGGCAATGGACCCCTAAGAACACCAGCTGCGACAGCGGCTTGGGGTGCCAGGA
CACGTTGATGCTCATTGAGAGCGGACCCCAAGTGAGCCTGGTGCTCTCCAAGGGCTGCACGG
AGGCCAAGGACCAGGAGCCCCGCGTCACTGAGCACCGGATGGGCCCCGGCCTCTCCCTGATC
TCCTACACCTTCGTGTGCCGCCAGGAGGACTTCTGCAACAACCTCGTTAACTCCCTCCCGCT
TTGGGCCCCACAGCCCCCAGCAGACCCAGGATCCTTGAGGTGCCAGTCTGCTTGTCTATGG
AAGGCTGTCTGGAGGGGACAACAGAAGAGATCTGCCCCAAGGGGACCACACACTGTTATGAT
GGCCTCCTCAGGCTCAGGGGAGGAGGCATCTTCTCCAATCTGAGAGTCCAGGGATGCATGCC
CCAGCCAGGTTGCAACCTGCTCAATGGGACACAGGAAATTGGGCCCCGTGGGTATGACTGAGA
ACTGCAATAGGAAAGATTTTCTGACCTGTCATCGGGGGACCACCATTATGACACACGGAAAC
TTGGCTCAAGAACCCACTGATTGGACCACATCGAATACCGAGATGTGCGAGGTGGGGCAGGT
GTGTCAGGAGACGCTGCTGCTCATAGATGTAGGACTCACATCAACCCTGGTGGGGACAAAAG
GCTGCAGCACTGTTGGGGCTCAAAATTCCCAGAAGACCACCATCCACTCAGCCCCCTCCTGGG
GTGCTTGTGGCCTCCTATACCCACTTCTGCTCCTCGGACCTGTGCAATAGTGCCAGCAGCAG
CAGCGTTCTGCTGAACTCCCTCCCTCCTCAAGCTGCCCCCTGTCCCAGGAGACCGGCAGTGTC
CTACCTGTGTGCAGCCCCCTTGGAACCTGTTCAAGTGGCTCCCCCGAATGACCTGCCCCAGG
GGCGCCACTCATTTGTTATGATGGGTACATTCACTCTCAGGAGGTGGGCTGTCCACCAAAAAT
GAGCATTCAGGGCTGCGTGGCCCAACCTTCCAGCTTCTTGTTGAACCACACCAGACAAATCG
GGATCTTCTCTGCGCGTGAGAAGCGTGATGTGCAGCCTCCTGCCTCTCAGCATGAGGGAGGT
GGGGCTGAGGGCCTGGAGTCTCTCACTTGGGGGGTGGGGCTGGCACTGGCCCCAGCGCTGTG
GTGGGGAGTGGTTTGCCCTTCTGCTTAACTCTATTACCCCCACGATTCTTCACCGCTGCTGA
CCACCCACACTCAACCTCCCTCTGACCTCATAACCTAATGGCCTTGGACACCAGATTCTTTC
CCATTCTGTCCATGAATCATCTTCCCCACACACAATCATTATATCTACTCACCTAACAGCA
ACACTGGGGAGAGCCTGGAGCATCCGGACTTGCCCTATGGGAGAGGGGACGCTGGAGGAGTG
GCTGCATGTATCTGATAATACAGACCCTGTCCTTTCA

FIGURE 10

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59847

><subunit 1 of 1, 437 aa, 1 stop

><MW: 46363, pI: 6.22, NX(S/T): 3

MSAVLLLALLGFILPLPGVQALLCQFGTVQHVWKVSDLPQWTPKNTSCDSGLGCQDTLMLI
ESGPQVSLVLSKGCTEAKDQEPRVTEHRMGPGLSLISYTFVCRQEDFCNNLVNSLPLWAPQP
PADPGSLRCPVCLSMEGCLEGTTEEICPKGTTHCYDGLLRLRGGGIFSNLRVQGCMPQPGCN
LLNGTQEIGPVGMTENCNRKDFLTCHRGTTIMTHGNLAQEPTDWTTSENTEMCEVGQVCQETL
LLIDVGLTSTLVGTKGCSTVGAQNSQKTTIHSAPPGVLVASYTHFCSSDLCNSASSSSVLLN
SLPPQAAPVPGDRQCPTCVQPLGTCSSGSPRMTCPRGATHCYDGYIHLSGGGLSTKMSIQGC
VAQPSSFLLNHTRQIGIFSAREKRDVQPPASQHEGGGAEGLESLTWGVGLALAPALWWGVVCPSC

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 243-260

N-glycosylation sites.

amino acids 46-49, 189-192, 382-385

Glycosaminoglycan attachment sites.

amino acids 51-54, 359-362

N-myristoylation sites.

amino acids 54-59, 75-80, 141-146, 154-159, 168-173, 169-174,
198-203, 254-259, 261-266, 269-274, 284-289, 333-338, 347-352,
360-365, 361-366, 388-393, 408-413, 419-424

FIGURE 11

CGACG**ATG**CTACGCGCGCCCGGCTGCCTCCTCCGGACCTCCGTAGCGCCTGCCGCGGCCCTG
GCTGCGGCGCTGCTCTCGTCGCTTGCGCGCTGCTCTTCTAGAGCCGAGGGACCCGGTGGC
CTCGTCGCTCAGCCCCCTATTTTCGGCACCAAGACTCGCTACGAGGATGTCAACCCCGTGCTAT
TGTCGGGCCCCGAGGCTCCGTGGCGGGACCTGAGCTGCTGGAGGGGACCTGCACCCCGGTG
CAGCTGGTCGCCCTCATTCGCCACGGCACCCGCTACCCACGGTCAAACAGATCCGCAAGCT
GAGGCAGCTGCACGGGTTGCTGCAGGCCCGCGGGTCCAGGGATGGCGGGGCTAGTAGTACCG
GCAGCCGCGACCTGGGTGCAGCGCTGGCCGACTGGCCTTTGTGGTACGCGGACTGGATGGAC
GGGCAGCTAGTAGAGAAGGGACGGCAGGATATGCGACAGCTGGCGCTGCGTCTGGCCTCGCT
CTTCCCGGCCCTTTTCAGCCGTGAGAACTACGGCCGCTGCGGCTCATCACCAGTTCGAAGC
ACCGCTGCATGGATAGCAGCGCCGCTTCTGCAGGGGCTGTGGCAGCACTACCACCCTGGC
TTGCCGCCCGCGGACGTGCGAGATATGGAGTTTGGACCTCCAACAGTTAATGATAAACTAAT
GAGATTTTTTGTACTGTGAGAAGTTTTTAACTGAAGTAGAAAAAATGCTACAGCTCTTT
ATCACGTGGAAGCCTTCAAACTGGACCAGAAATGCAGAACATTTTAAAAAAGTTGCAGCT
ACTTTGCAAGTGCCAGTAAATGATTTAAATGCAGATTTAATTCAAGTAGCCTTTTTTCACCTG
TTCATTTGACCTGGCAATTAAAGGTGTTAAATCTCCTTGGTGTGATGTTTTTGACATAGATG
ATGCAAAGGTATTAGAATATTTAAATGATCTGAAACAATATTGGAAAAGAGGATATGGGTAT
ACTATTAACAGTCGATCCAGCTGCACCTTGTTTCAGGATATCTTTCAGCACTTGGACAAAGC
AGTTGAACAGAAACAAAGGTCTCAGCCAATTTCTTCTCCAGTCATCCTCCAGTTTGGTCATG
CAGAGACTCTTCTTCCACTGCTTTCTCTCATGGGCTACTTCAAAGACAAGGAACCCCTAACA
GCGTACAATTACAAAAACAAATGCATCGGAAGTTCCGAAGTGGTCTCATTGTACCTTATGC
CTCGAACCTGATATTTGTGCTTTACCACTGTGAAAATGCTAAGACTCCTAAAGAACAAATTCC
GAGTGCAGATGTTATTAAATGAAAAGGTGTTACCTTTGGCTTACTCACAAGAACTGTTTCA
TTTTATGAAGATCTGAAGAACCACTACAAGGACATCCTTCAGAGTTGTCAAACCAGTGAAGA
ATGTGAATTAGCAAGGGCTAACAGTACATCTGATGAACTA**TGA**GTAAGTGAAGAACATTTTT
AATTCTTTAGGAATCTGCAATGAGTGATTACATGCTTGTAATAGGTAGGCAATTCCTTGATT
ACAGGAAGCTTTTATATTACTTGAGTATTTCTGTCCTTTTCACAGAAAAACATTGGGTTTCTC
TCTGGGTTTGGACATGAAATGTAAGAAAAGATTTTTTCACTGGAGCAGCTCTCTTAAGGAGAA
ACAAATCTATTTAGAGAAACAGCTGGCCCTGCAAATGTTTACAGAAATGAAATTCTTCCTAC
TTATATAAGAAATCTCACACTGAGATAGAATTGTGATTTTCATAATAACACTTGAAAAGTGCT
GGAGTAACAAAATATCTCAGTTGGACCATCCTTAACCTTGATTGAACTGTCTAGGAACCTTAC
AGATTGTTCTGCAGTTCTCTCTTTTCTCCTCAGGTAGGACAGCTCTAGCATTTTCTTAATC
AGGAATATTGTGGTAAGCTGGGAGTATCACTCTGGAAGAAAGTAACATCTCCAGATGAGAAT
TTGAAACAAGAAACAGAGTGTTGTAAGGACACCTTCACTGAAGCAAGTCGGAAGGTACAA
TGAAAATAAATATTTTTGGTATTTATTTATGAAATATTTGAACATTTTTTCAATAATTCCTT
TTTACTTCTAGGAAGTCTCAAAGACCATCTTAAATATTATATGTTTGGACAATTAGCAAC
AAGTCAGATAGTTAGAATCGAAGTTTTTCAAATCCATTGCTTAGCTAACTTTTTTCATTCTGT
CACTTGGCTTCGATTTTTATATTTTCTATTTATGAAATGTATCTTTTGGTTGTTTGATTT
TTCTTTCTTTCTTTGTAAATAGTTCTGAGTTCTGTCAAATGCCGTGAAAGTATTTGCTATAA
TAAAGAAAATTCTTGTGACTTTAAAAAAA

[illegible]

```
><subunit 1 of 1, 487 aa, 1 stop
```

MLRAPGCLLRTSVAPAAALAAALLSSLARCSLLEPRDPVASSLSPTYFGTKTRYEDVNPVLLS
GPEAPWRDPELLEGTCTPVQLVALIRHGTRYPTVKQIRKLRQLHGLLQARGSRDGGASSTGS
RDLGAALADWPLWYADWMDGQLVEKGRQDMRQLALRLASLFPALFSRENYGRLRLITSSKHR
CMDSSAAFLQGLWQHYHPGLPPPDVADMEFGPPTVNDKLMRFFDHCEKFLTEVEKNATALYH
VEAFKTGPEMQNILKKVAATLQVPVNDLNADLIQVAFFTC SFDLAIKGVKSPWCDVFDIDDA
KVLEYLNDLQYWKRGYGYTINSRSSCTLFQDIFQHLDKAVEQKQRSQPISSPVILQFGHAE
TLLPLLSLMGYFKDKEPLTAYNYKKQMRKFRSGLIVPYASNLI FVLYHCENAKTPKEQFRV
QMLLNEKVLPLAYS OETVSFYEDLKNHYKDILOSCOTSEECELARANSTSD

Signal sequence

N-glycosylation sites.

N-myristoylation sites.

amino acids 107-113, 113-119, 117-123, 118-124, 128-134

Endoplasmic reticulum targeting sequence.

amino acids 484-489

FIGURE 13

GGGACTACAAGCCGCGCCGCGCTGCCGCTGGCCCCCTCAGCAACCCTCGACATGGCGCTGAGGCGGCCACCGCGAC
TCCGGCTCTGCGCTCGGCTGCCTGACTTCTTCCTGCTGCTGCTTTTCAGGGGCTGCCTGATAGGGGCTGTAAATC
TCAAATCCAGCAATCGAACCCAGTGGTACAGGAATTTGAAAGTGTGGAACGTGTCTTGCATCATTACGGATTTCGC
AGACAAGTGAACCCAGGATCGAGTGAAGAAAATTCAAGATGAACAAACCACATATGTGTTTTTTTGACAACAAAA
TTCAGGGAGACTTGGCGGGTCGTGCAGAAATACTGGGGAAGACATCCCTGAAGATCTGGAATGTGACACGGAGAG
ACTCAGCCCTTTATCGCTGTGAGGTCTGTGCTCGAAATGACCGCAAGGAAATTGATGAGATTGTGATCGAGTTAA
CTGTGCAAGTGAAGCCAGTGACCCCTGTCTGTAGAGTGCCGAAGGCTGTACCAGTAGGCAAGATGGCAACACTGC
ACTGCCAGGAGAGTGAGGGCCACCCCGGCCTCACTACAGCTGGTATCGCAATGATGTACCACTGCCACGGATT
CCAGAGCCAATCCCAGATTTTCGCAATTCTTCTTTCCACTTAAACTCTGAAACAGGCACCTTTGGTGTTCCTGCTG
TTCACAAGGACGACTCTGGGCAGTACTACTGCATTGCTTCCAATGACGCAGGCTCAGCCAGGTGTGAGGAGCAGG
AGATGGAAGTCTATGACCTGAACATTGGCGGAATTATTGGGGGGTCTGGTTGTCTTGTCTGTACTGGCCCTGA
TCACGTTGGGCATCTGCTGTGCATACAGACGTGGCTACTTCATCAACAATAAACAGGATGGAGAAAGTTACAAGA
ACCCAGGGAACAGATGGAGTTAACTACATCCGCACTGACGAGGAGGGCGACTTCAGACACAAGTCATCGTTTG
TGATCTGAGACCCGCGGTGTGGCTGAGAGCGCACAGAGCGCACGTGCACATACCTCTGCTAGAAACTCCTGTCAA
GGCAGCGAGAGCTGATGCACTCGGACAGAGCTAGACACTCATTGAGAAGCTTTTCGTTTTTGGCCAAAGTTGACCA
CTACTCTTCTTACTCTAACAAGCCACATGAATAGAAGAATTTTCCTCAAGATGGACCCGGTAAATATAACCACAA
GGAAGCGAAACTGGGTGCGTTCACTGAGTTGGGTTTCTAATCTGTTTCTGGCCTGATTCGCCGATGAGTATTAGG
GTGATCTTAAAGAGTTTGCTCACGTAAACGCCCCGTGCTGGGCCCTGTGAAGCCAGCATGTTCAACCACTGGTCGTT
CAGCAGCCACGACAGCACCATGTGAGATGGCGAGGTGGCTGGACAGCACCAGCAGCGCATCCCGCGGGAAACCCA
GAAAAGGCTTCTTACACAGCAGCCTTACTTCATCGGCCCACAGACACCACCGCAGTTTCTTCTTAAAGGCTCTGC
TGATCGGTGTGTCAGTGTCCATTGTGGAGAAGCTTTTGGATCAGCATTGTGTAACCAACCAAAATCAGGAAG
GTAAATTGGTTGCTGGAAGAGGGATCTTGCCTGAGGAACCTGCTTGTCCAACAGGGTGTGAGGATTTAAGGAAA
ACCTTCGTCTTAGGCTAAGTCTGAAATGGTACTGAAATATGCTTTTCTATGGGTCTTGTGTTATTTTATAAAATTT
TACATCTAAATTTTGTCTAAGGATGTATTTTGATTATTGAAAAGAAAATTTCTATTTAAACTGTAAATATATTGT
CATACAATGTAAATAACCTATTTTTTTTAAAAAAGTTCAACTTAAGGTAGAAGTTCCAAGCTACTAGTGTAAAT
TGAAAAATATCAATAATTAAGAGTATTTTACCCAAGGAATCCTCTCATGGAAGTTTACTGTGATGTTTCTTTCT
CACACAAGTTTTAGCCTTTTTTACAAGGGAACCTACTGTCTACACATCAGACCATAGTTGCTTAGGAAACCTT
TAAAAATTCAGTTAAGCAATGTTGAAATCAGTTTGCATCTCTTCAAAAGAAACCTCTCAGGTTAGCTTTGAACT
GCCTCTTCTGAGATGACTAGGACAGTCTGTACCCAGAGGCCACCCAGAAGCCCTCAGATGTACATACACAGATG
CCAGTCAGCTCCTGGGGTGTGCGCCAGGCGCCCCCGCTCTAGCTCACTGTTGCCTCGCTGTCTGCCAGGAGGCCCT
GCCATCCTTGGGCCCCTGGCAGTGGCTGTGTCCAGTGAGCTTTACTCACGTGGCCCTTGCTTCATCCAGCACAGC
TCTCAGGTGGGCACTGCAGGGACACTGGTGTCTTCCATGTAGCGTCCCAGCTTTGGGCTCCTGTAACAGACCTCT
TTTTGGTTATGGATGGCTCACAAAATAGGGCCCCCAATGCTATTTTTTTTTTTAAGTTTGTGTTAATTATTTGTT
AAGATTGTCTAAGGCCAAAGGCAATTGCGAAATCAAGTCTGTCAAGTACAATAACATTTTTTAAAGAAAATGGAT
CCCCTGTTTCTCTTTGCCACAGAGAAAGCACCCAGACGCCACAGGCTCTGTGCGATTTCAAAACAAACCATGAT
GGAGTGGCGGCCAGTCCAGCCTTTTAAAGAACGTGAGGTGGAGCAGCCAGGTGAAAGGCCCTGGCGGGGAGGAAAG
TGAAACGCCTGAATCAAAAGCAGTTTTCTAATTTTGACTTTAAATTTTTTATCCGCCGAGACACTGCTCCCAT
TGTGGGGGGACATTAGCAACATCACTCAGAAGCCTGTGTTCTTCAAGAGCAGGTGTTCTCAGCCTCATATGCCCT
GCCGTGCTGGACTCAGGACTGAAGTGCTGTAAAGCAAGGAGCTGCTGAGAAGGAGCACTCCACTGTGTGCCCTGGA
GAATGGCTCTCACTACTACCTTGTCTTTTACGCTTCCAGTGTCTTGGGTTTTTATACCTTTGACAGCTTTTTTTT
AATTGCATACATGAGACTGTGTTGACTTTTTTTTAGTTATGTGAAACACTTTGCCGAGGCCGCTGGCAGAGGCA
GGAAATGCTCCAGCAGTGGCTCAGTGCTCCCTGGTGTCTGCTGCATGGCATCCTGGATGCTTAGCATGCAAGTTC
CCTCCATCATTTGCCACCTTGGTAGAGAGGGATGGCTCCCCACCCTCAGCGTTGGGGATTACAGCTCCAGCCTCCT
TCTTGGTTGTCTAGTGATAGGGTAGCCTTATTGCCCCCTCTTCTTATACCCTAAAACCTTCTACACTAGTGCCA
TGGGAACCAAGGTCTGAAAAAGTAGAGAGAAGTGAAAGTAGAGTCTGGGAAGTAGCTGCCTATAACTGAGACTAGA
CGGAAAAGGAATACTCGTGTATTTTAAGATATGAATGTGACTCAAGACTCGAGGCCGATACGAGGCTGTGATTCT
GCCTTTGGATGGATGTTGCTGTACACAGATGCTACAGACTTGTAACACACCGTAATTTGGCATTTGTTTAAC
CTCATTATATAAAGCTTCAAAAAAACCCA

FIGURE 14

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77624

><subunit 1 of 1, 310 aa, 1 stop

><MW: 35020, pI: 7.90, NX(S/T): 3

MALRRPPRLRLCARLPDFFLLLLFRGCLIGAVNLKSSNRTPVVQEFESVELSCIITDSQTS
PRIEWKKIQDEQTTYVFFDNKIQGDLAGRAEILGKTSCLKIWNVTRRDSALYRCEVVARNDRK
EIDEIVIELTVQVKPVTVPVCRVPKAVPVGKMATLHCQSEGHPRPHYSWYRNDVPLPTDSRA
NPRFRNSSFHLNSETGTLVFTAVHKDDSGQYYCIASNDAGSARCEEQEMEVDNLNIGGIIGG
VLVVLAVLALITLGICCAYYRRGYFINNKQDGESYKNPGKPDGVNYIRTDEEGDFRHKSSFVI

Important features of the protein:

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 243-263

N-glycosylation sites.

amino acids 104-107, 192-195

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 107-110

Casein kinase II phosphorylation site.

amino acids 106-109, 296-299

Tyrosine kinase phosphorylation site.

amino acids 69-77

N-myristoylation sites.

amino acids 26-31, 215-220, 226-231, 243-248, 244-249, 262-267

0662001

FIGURE 15

CAGGACCAGGTCTTCTACGCTGGAGCAGCGGGGAGACAGCCACCATGCACATCCTCGTGGTCCATGCCATGGTG
ATCCTGCTGACGCTGGGCCCCGCTCGAGCCGACGACAGCGAGTTCCAGGCGCTGCTGGACATCTGGTTTCCGGAG
GAGAAGCCACTGCCCACCGCCTTCTGGTGGACACATCGGAGGAGGCGCTGCTGCTTCCCTGACTGGCTGAAGCTG
CGCATGATCCGTTCTGAGGTGCTCCGCCTGGTGGACGCCGCCCTGCAGGACCTGGAGCCGCAGCAGCTGCTGCTG
TTCGTGCAGTCGTTTGGCATCCCCGTGTCCAGCATGAGCAAACCTCCTCCAGTTCCCTGGACCAGGCAGTGCCCCAC
GACCCCCAGACTCTGGAGCAGAACATCATGGACAAGAATTACATGGCCCACCTGGTGGAGGTCCAGCATGAGCGC
GGCGCCTCCGGAGGCCAGACTTTCCACTCCTTGCTCACAGCCTCCCTGCCGCCCCGCCGAGACAGCACAGAGGCA
CCCAAACCAAAGAGCAGCCCAGAGCAGCCCATAGGCCAGGGCCGGATTCCGGTGGGGACCCAGCTCCGGGTGCTG
GGCCCTGAGGACGACCTGGCTGGCATGTTTCTCCAGATTTTCCCGCTCAGCCCGGACCCTCGGTGGCAGAGCTCC
AGTCCCCGCCCGCTGGCCCTCGCCCTGCAGCAGGCCCTGGGCCAGGAGCTGGCCCGCGTCTCCAGGGCAGCCCC
GAGGTGCCGGGCATCACGGTGCCTGCTGTCAGGCCCTCGCCACCCTGCTCAGCTCCCCACACGGCGGTGCCCTG
GTGATGTCCATGCACCGTAGCCACTTCTGGCCTGCCCGCTGCTGCGCCAGCTCTGCCAGTACCAGCGTGTGTG
CCACAGGACACCGGCTTCTCCTCGCTCTTCTGAAGGTGCTCCTGCAGATGCTGCAGTGGCTGGACAGCCCTGGC
GTGGAGGGCGGGCCCCCTGCCGGGCACAGCTCAGGATGCTTGCCAGCCAGGCCCTCAGCCGGGCGCAGGCTCAGTGAT
GTGCGAGGGGGGGCTCCTGCGCCTGGCCGAGGCCCTGGCCTTCCGTGACGACCTGGAGGTGGTCAGCTCCACCGTC
CGTGCCGTCATCGCCACCCTGAGGTCTGGGGAGCAGTGACGCTGGAGCCGGACCTGATCAGCAAAGTCTCCAG
GGGCTGATCGAGGTGAGGTCCCCCACCTGGAGGAGCTGCTGACTGCATTCTTCTCTGCCACTGCGGATGCTGCC
TCCCCGTTTCCAGCCTGTAAGCCCCGTTGTGGTGGTGAGCTCCCTGCTGCTGCAGGAGGAGGAGCCCTGGCTGGG
GGGAAGCCGGGTGCCGACGGTGGCAGCCTGGAGGCCGTGCGGCTGGGGCCCTCGTCAGGCCCTCTAGTGGACTGG
CTGGAATGCTGGACCCCCAGGTGGTCAGCAGTCCCCGACCTGCAGCTCAGGCTGCTCTTCTCCCGAGGAAG
GGCAAAGGTCAGGCCCAGGTGCCCTCGTTCCGCTCCCTACCTCCTGACCCTTTCACGCATCAGTCCAGCTGGCCCC
ACACTGCACCAAGTGATCCGAGTCTTGCTGGGCAAGAGCCGGGAACAGAGGTTTCGACCCCTCTGCCTCTCTGGAC
TTCTCTGGGCCTGCATCCATGTTCTCGCATCTGGCAGGGGCGGGACCAGCGCACCCCGCAGAAGCGGCGGGAG
GAGCTGGTGTGTCGGGTCCAGGGCCCCGAGCTCATCAGCCTGGTGGAGCTGATCCTGGCCGAGGCGGAGACGCGG
AGCCAGGACGGGGACACAGCCGCTGCAGCCTCATCCAGGCCCGGCTGCCCTGCTGCTCAGCTGCTGCTGTGGG
GACGATGAGAGTGTGAGGAAGGTGACGGAGCAGCTGTCAGGCTGCATCCAGCAGTGGGGAGACAGCGTGTGGGA
AGGCGCTGCCGAGACCTTCTCCTGCAGCTCTACCTACAGCGGCCGGAGCTGCGGGTGCCCGTGCCTGAGGTCTTA
CTGCACAGCGAAGGGGCTGCCAGCAGCAGCTGCTGCAAGCTGGACGGACTCATCCACCGCTTCATCACGCTCCTT
GCGGACACCAGCAGTCCCCGGGCGTTGGAGAACCAGGGGCGGATGCCAGCATGGCCTGCCGGAAGCTGGCGGTG
GCGCACCCGCTGCTGCTGCTCAGGCACCTGCCCATGATCGCGGCGCTCCTGCACGGCCGCACCCACCTCAACTTC
CAGGAGTTCGGGCAGCAGAACCACCTGAGCTGCTTCTGTCAGTGTGCGGCTGCTGGAGCTGCTGCAGCCGCAC
GTGTTCCGCAGCGAGCACCAGGGGGCGCTGTGGGACTGCCTTCTGTCTTCATCCGCTGCTGCTGAATTACAGG
AAGTCTTCCCGCCATCTGGCTGCCTTCATCAACAAGTTTGTGAGTTTCATCCATAAGTACATTACCTACAATGCC
CCAGCAGCCATCTCCTTCTGTCAGAAGCACGCCGACCCGCTCCACGACCTGTCTTCGACAACAGTGACCTGGTG
ATGCTGAAATCCCTCCTTGCAGGGCTCAGCCTGCCCAGCAGGGACGACAGGACCGACCGAGGCCCTGGACGAAGAG
GGCGAGGAGGAGAGCTCAGCCGGCTCCTTGCCCTGGTCAGCGTCTCCTGTTACCCCTCTGACCGCGGCCGAG
ATGGCCCCCTACATGAAACGGCTTTCCCGGGGCCAAACGGTGGAGGATCTGCTGGAGGTTCTGAGTGACATAGAC
GAGATGTCCCGCGGAGACCCGAGATCCTGAGCTTCTTCTCGACCAACCTGCAGCGGCTGATGAGCTCGGCCGAG
GAGTGTGTCGCGCAACCTCGCCTTCAGCCTGGCCCTGCGCTCCATGCAGAACAGCCCCAGCATTCAGCCGCTTTC
CTGCCCACGTTTCATGTACTGCCTGGGCAGCCAGGACTTTGAGGTGGTGCAGACGGCCCTCCGGAACCTGCCTGAG
TACGCTCTCCTGTGCCAAGAGCACGCGGCTGTGCTGCTCCACCGGGCCTTCTGGTGGGCATGTACGGCCAGATG
GACCCACAGCGCGCAGATCTCCGAGGCCCTGAGGATCCTGCATATGGAGGCCGTGATGTGAGCCTGTGGCAGCCGA
CCCCCTCCAAGCCCCGGCCCGTCCCGTCCCCGGGGATCCTCGAGGCAAAGCCCAGGAAGCGTGGGCGTTGCTGG
TCTGTCCGAGGAGGTGAGGGCGCCGAGCCCTGAGGCCAGGCAGGCCAGGAGCAATACTCCGAGCCCTGGGGTGG
CTCCGGGCGGCGGCTGGCATCAGGGGCCGTCCAGCAAGCCCTCATTACCTTCTGGGCCACAGCCCTGCCGCGG
AGCGGCGGATCCCCCGGGCATGGCCTGGGCTGGTTTTGAATGAAACGACCTGAAGTGTCAA

[illegible]

><MW: 114213, pI: 6.42, NX(S/T): 0

amino acids 329-333, 634-638

FIGURE 17

CCGGGCCATGCAGCCTCGGCCCCGCGGGCGCCCGCCGCGCACCCGAGGAGATGAGGCTCCGC
AATGGCACCTTCCTGACGCTGCTGCTCTTCTGCCTGTGCGCCTTCCTCTCGCTGTCCTGGTA
CGCGGCACTCAGCGGCCAGAAAGGCGACGTTGTGGACGTTTACCAGCGGGAGTTCCTGGCGC
TGCGCGATCGGTTGCACGCAGCTGAGCAGGAGAGCCTCAAGCGCTCCAAGGAGCTCAACCTG
GTGCTGGACGAGATCAAGAGGGCCGTGTGAGAAAGGCAGGCGCTGCGAGACGGAGACGGCAA
TCGCACCTGGGGCCGCTAACAGAGGACCCCCGATTGAAGCCGTGGAACGGCTCACACCGGC
ACGTGCTGCACCTGCCCACCGTCTTCCATCACCTGCCACACCTGCTGGCCAAGGAGAGCAGT
CTGCAGCCCCGCGGTGCGCGTGGGCCAGGGCCGCACCGGAGTGTCGGTGGTGATGGGCATCCC
GAGCGTGCGGCGGAGGTGCACTCGTACCTGACTGACACTCTGCACTCGCTCATCTCCGAGC
TGAGCCCGCAGGAGAAGGAGGACTCGGTCACTGCTGGTGCTGATCGCCGAGACTGACTCACAG
TACACTTCGGCAGTGACAGAGAACATCAAGGCCTTGTTCCCCACGGAGATCCATTCTGGGCT
CCTGGAGGTCATCTCACCTCCCCCACTTCTACCCTGACTTCTCCCGCCTCCGAGAGTCCT
TTGGGGACCCCAAGGAGAGAGTCAGGTGGAGGACCAAACAGAACCTCGATTACTGCTTCCTC
ATGATGTACGCGCAGTCCAAAGGCATCTACTACGTGCAGCTGGAGGATGACATCGTGGCCAA
GCCCAACTACCTGAGCACCATGAAGAACTTTGCACTGCAGCAGCCTTCAGAGGACTGGATGA
TCCTGGAGTTCTCCAGCTGGGCTTCATTGGTAAGATGTTCAAGTCGCTGGACCTGAGCCTG
ATTGTAGAGTTCATTCTCATGTTCTACCGGGACAAGCCCATCGACTGGCTCCTGGACCATAT
TCTGTGGGTGAAAGTCTGCAACCCCGAGAAGGATGCGAAGCACTGTGACCGGCAGAAAGCCA
ACCTGCGGATCCGCTTCAAACCGTCCCTCTTCCAGCACGTGGGCACTCACTCCTCGCTGGCT
GGCAAGATCCAGAACTGAAGGACAAAGACTTTGGAAAGCAGGCGCTGCGGAAGGAGCATGT
GAACCCGCCAGCAGAGGTGAGCACGAGCCTGAAGACATACCAGCACTTCACCCTGGAGAAAG
CCTACCTGCGCGAGGACTTCTTCTGGGCCTTACCCCTGCCGCGGGGGACTTCATCCGCTTC
CGCTTCTTCCAACCTCTAAGACTGGAGCGGTTCTTCTTCCGCAGTGGAACATCGAGCACCC
GGAGGACAAGCTCTTCAACACGTCTGTGGAGGTGCTGCCCTTCGACAACCCTCAGTCAGACA
AGGAGGCCCTGCAGGAGGGCCGCACCGCCACCCTCCGGTACCCTCGGAGCCCCGACGGCTAC
CTCCAGATCGGCTCCTTCTACAAGGGAGTGGCAGAGGGAGAGGTGGACCCAGCCTTCGGCCC
TCTGGAAGCACTGCGCCTCTCGATCCAGACGGAATCCCCTGTGTGGGTGATTCTGAGCGAGA
TCTTCTGAAAAAGGCCGACTAAGCTGCGGGCTTCTGAGGGTACCCTGTGGCCAGCCCTGAA
GCCACATTTCTGGGGGTGTCGTCACTGCCGTCCCCGGAGGGCCAGATACGGCCCCGCCCAA
AGGGTTCTGCCTGGCGTGGGGCTTGGGCCGGCCTGGGGTCCGCCGCTGGCCCCGAGGCCCTA
GGAGCTGGTGCTGCCCCCGCCCCGCCGGGCCGCGGAGGAGGCAGGCGGGCCCCCACACTGTGCC
TGAGGCCCGGAACCGTTTCGCACCCGGCCTGCCCCAGTCAGGCCGTTTTAGAAAGAGCTTTTAC
TTGGGCGCCCCCGCTCTCTGGCGCGAACACTGGAATGCATATACTACTTTATGTGCTGTGTT
TTTTATTCTTGATACATTTGATTTTTTTCACGTAAGTCCACATATACTTCTATAAGAGCGTG
ACTTGTAATAAAGGGTTAATGAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAA

FIGURE 18

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA82307

><subunit 1 of 1, 548 aa, 1 stop

><MW: 63198, pI: 8.10, NX(S/T): 4

MRLRNGTFLTLLLFCLCAFLSLSWYAALSGQKGDVVDVYQREFLALRDLHAAEQESLKRSK
ELNLVLDEIKRAVSERQALRDGDGNRTWGRLTEDPRLKPWNGSHRHVLHLPTVFFHLLPHLLA
KESSLQPAVRVGQGRGTGVSVMGIPSVRREVHSYLTDTLHSLISELSPQEKEDSVIVVLIAE
TDSQYTSAVTENIKALFPTEIHSGLLEVISPSPHFYPDFSRLRESFGDPKERVVRWRTKQNL
YCFLMMYAQSKGIYYVQLEDDIVAKPNYLSTMKNFALQQPSEDWMILEFSQLGFIGKMFKSL
DLSSLIVEFILMFYRDKPIDWLLDHILWVKVCNPEKDAKHCDRQKANLRIRFKPSLFQHVGT
SSLAGKIQKLKDKDFGKQALRKEHVNPPAEVSTSLKTYQHFTLEKAYLREDFFWAFTPAAGD
FIRFRFFQPLRLERFFFRSGNIEHPEDKLFNTSVEVLPPDNPQSDKEALQEGRTATLRYPRS
PDGYLQIGSFYKGVAEGEVDPAFGPLEALRLSIQTDSPVWVILSEIFLKKAD

Important features:

Signal sequence

amino acids 1-23

N-glycosylation sites.

amino acids 5-9, 87-91, 103-107, 465-469

N-myristoylation sites.

amino acids 6-12, 136-142, 370-376, 509-515